S8 IMMUNOTOXICITY STUDIES FOR HUMAN PHARMACEUTICALS

For questions regarding this draft document contact (CDER) Kenneth L. Hastings at 301-827-5922.

IMMUNOTOXICITY STUDIES FOR HUMAN PHARMACEUTICALS

ICH Harmonised Tripartite Guideline

Having reached *Step 2* of the ICH Process at the ICH Steering Committee meeting on 18 November 2004, the Committee releases this document for consultation.

TABLE OF CONTENTS

1. INTRODUCTION

Evaluation of potential adverse effects of human pharmaceuticals on the immune system should be incorporated into standard drug development. Toxicity to the immune system encompasses a variety of adverse effects. These include suppression or enhancement of the immune response. Suppression of the immune response can lead to decreased host resistance to infectious agents or tumor cells, whereas enhancing the immune response can stimulate the expansion of autoreactive immune cells and lead to autoimmune diseases. Drug or drug-protein adducts might also be recognized as foreign and stimulate an anti-drug response. Subsequent exposures to the drug can lead to hypersensitivity (allergic) reactions. Much of the science and method development and validation efforts in the past have been focused on evaluating drug development candidates for their potential to be either immunosuppressive or contact (dermal) sensitizers.

1.1 Objectives of the guideline

The objectives of this guideline are to provide (1) recommendations on nonclinical testing approaches to identify compounds which have the potential to be immunosuppressive, and (2) guidance on a weight-of-evidence decision making approach for immunotoxicity testing.

1.2 Background

The most robust methods available for practical use in the assessment of adverse drug effects on immune function are those designed to detect and evaluate immunosuppression. Historically, unintended immunosuppression has been causally related to antiproliferative drugs primarily used to treat cancer. In such instances, adverse findings in nonclinical studies are predictive of human immunotoxicity in a rather straightforward manner. That is, specific assays to determine immunotoxicity are probably not valuable in drug risk assessment since the target tissues are usually rapidly dividing cell types, such as bone marrow-derived immune system progenitor cells. Hence, the adverse effects on immune function can be predicted based on pharmacologic activity and can usually be reliably modelled in animals.

It has become apparent in recent years that immunosuppression can be associated with other types of drugs. It is possible to divide these drugs into two distinct groups: Those that are intended to modulate immune function for therapeutic purposes (e.g. drugs intended to prevent organ transplant rejection) where adverse immunosuppression can be considered exaggerated pharmacodynamics, and those that are not intended to affect immune function but can cause unintended immunosuppression due for instance, to induced necrosis, apoptosis of immune cells or interaction with cellular receptors shared by both target tissues and non-target immune system cells. Although this difference is relatively obvious, distinction between exaggerated pharmacodynamics and non-target toxicity can be less obvious for certain classes of drugs (e.g. anti-inflammatory drugs).

1.3 Scope of the Guideline

This guideline is focused on providing recommendations on nonclinical testing for immunosuppression induced by low molecular weight drugs (non-biologicals). This

guideline applies to new pharmaceuticals intended for use in humans, as well as to marketed drug products proposed for different indications or other variations on the current product label in which the change could result in unaddressed and relevant toxicologic issues. In addition, the guideline might also apply to drugs in which clinical signs of immunosuppression are observed during clinical trials and following approval to market. The term immunotoxicity in this guideline will primarily refer to immunosuppression, i.e. a state of increased susceptibility to infections or the development of tumors.

It is beyond the scope of this guideline to provide specific guidance on how each immunotoxicity study should be performed. General guidance is provided in Appendix 1.

1.4 Overview

The general principles that apply to this guideline are:

1) All new investigational drugs should be evaluated for the potential to produce immunosuppression.

2) Methods include standard toxicity studies (STS) and additional immunotoxicity studies conducted as appropriate. Whether additional immunotoxicity studies are appropriate should be determined by a weight of evidence review of cause(s) for concern.

The description of the guideline below will follow the flow diagram shown in Figure 1. More detailed descriptions of the testing methods are described in Appendix 1.

2. GUIDELINE

2.1 Assessment of potential immunotoxicity

The initial screen for potential immunotoxicity involves the standard toxicity studies. Data from rodent and non-rodent studies from early short term to more chronic repeat-dose studies should be taken into consideration. Additional details on the parameters that should be evaluated and the reporting of histopathology findings are provided in Appendix 1.

In addition to the findings from the STS, other causes for concern that might prompt additional immunotoxicity studies include: (1) the pharmacological properties of the drug, (2) the intended patient population, (3) known drug class effects, and (4) the disposition of the drug.

The flow diagram (Figure 1) illustrates the recommended decision process in immunotoxicity evaluation.

2.1.1 Standard toxicity studies

Data from STS should be evaluated for signs of immunotoxic potential. Signs that should be taken into consideration are the following.

(1) Hematological changes - Evidence of myelosuppression, usually seen in peripheral blood changes (e.g. pancytopenia, leukopenia, lymphopenia, or other blood dyscrasias);
(2) Alterations in immune system organ weights and histology (e.g. changes in thymus, spleen, lymph nodes, and/or bone marrow);
(3) Decreased basal serum immunoglobulins – serum globulins are a rather insensitive marker of immunotoxicity due to the long half life of immunoglobulins. However, changes in globulins that occur without a plausible explanation can indicate potential immunotoxicity.
(4) Increased incidence of infections.
(5) Evidence of carcinogenicity, especially in the absence of genotoxicity.

If the findings from the STS indicate that there are signs of immunotoxicity, the decision to conduct additional immunotoxicity testing should be considered in a weight-of-evidence review of the data. Similar to the assessment of risk with toxicities in other organ systems, the assessment of immunotoxicity should include the following:

- statistical and biological significance of the changes,
- severity of the effects,
- dose dependency,
- safety factor above the expected clinical dose,
- study duration,
- number of species and endpoints affected,
- changes that may occur secondarily to other factors (eg. stress, see appendix 1),
- possible cellular targets and/or mechanism of action,
- doses which produces these changes in relation to doses which produce other toxicities and
- reversibility of effect(s).

2.1.2 Other Causes for Concern in the Weight-of-Evidence Review

The following factors should also be considered:

(1) If the pharmacological properties of a test compound indicate it has the potential to produce significant immunosuppression, additional immunotoxicity testing should be considered. For example, anti-inflammatory drugs are known to affect the function of certain types of cells of the immune system; however, their ability to suppress adaptive and/or innate immune responses is not clear. Information on the ability of the compound to affect the immune system can be gathered as part of the pharmacological studies conducted during the discovery or early development phases. These non-GLP pharmacology studies could be used in deciding if additional immunotoxicity studies are needed. The decision to conduct additional immunotoxicity studies should be based on a weight of evidence approach.

(2) The targeted patient population should also be considered. For instance, additional immunotoxicity testing might be needed if the majority of the targeted patient population is immunocompromised.

(3) Compounds structurally similar to compounds with known immunosuppressive properties should also be considered for additional immunotoxicity testing.

(4) If the compound and/or its metabolites are known to be retained at high concentrations in cells of the immune system, additional immunotoxicity testing should be considered.

If signs of immunotoxicity are observed in STS and/or one of the above four factors apply, it is recommended that the sponsor conduct studies of drug effect on immune function or provide justification for not performing these evaluations.

2.2 Selection and Design of Additional Immunotoxicity Studies

2.2.1 Selection of assays

If the weight-of-evidence approach indicates that additional immunotoxicity studies are needed, there are a number of animal models which can be used. If there are changes in standard toxicity testing data suggesting immunosuppression, the type of additional immunotoxicity testing that is appropriate will depend on the nature of the immunological changes observed and concerns raised by the class of compound.

It is recommended that an immune function study be conducted. Where a specific target is not identified, an immune function study such as the T-cell dependent antibody response (TDAR) is recommended. If specific cell types are affected in STS, assays that measure function of that specific cell type might be conducted (see appendix 1).

Immunophenotyping of leukocyte populations, a non-functional assay, can be conducted to identify the specific cell populations affected and useful clinical biomarkers.

2.2.2 Study Design

It is a generally accepted study design to assess drug-induced immunosuppression in studies with 28 consecutive daily oral doses in mice or rats. The species, dose, duration, and route of administration used in immune function assays should be consistent, where possible, with the nonclinical toxicology study in which an adverse immune effect was observed. The high dose should be above the no observed adverse effect level (NOAEL) but below a level inducing changes secondary to stress. Multiple dose levels are recommended in order to determine dose-response relationships and the dose at which no immunotoxicity is observed. Adaptations of immune function assays developed in rodents have been described using non-rodent species. Under most circumstances, immunological test methods can be appropriately modified for these other species.

2.2.3 Evaluation of Need for Follow-up Testing

Results of the entire data set should be evaluated as to whether sufficient data are available to reasonably determine the risk of immunotoxicity. If the overall risk-benefit analysis suggests that the risk of immunotoxicity is acceptable, then no follow-up testing might be called for.

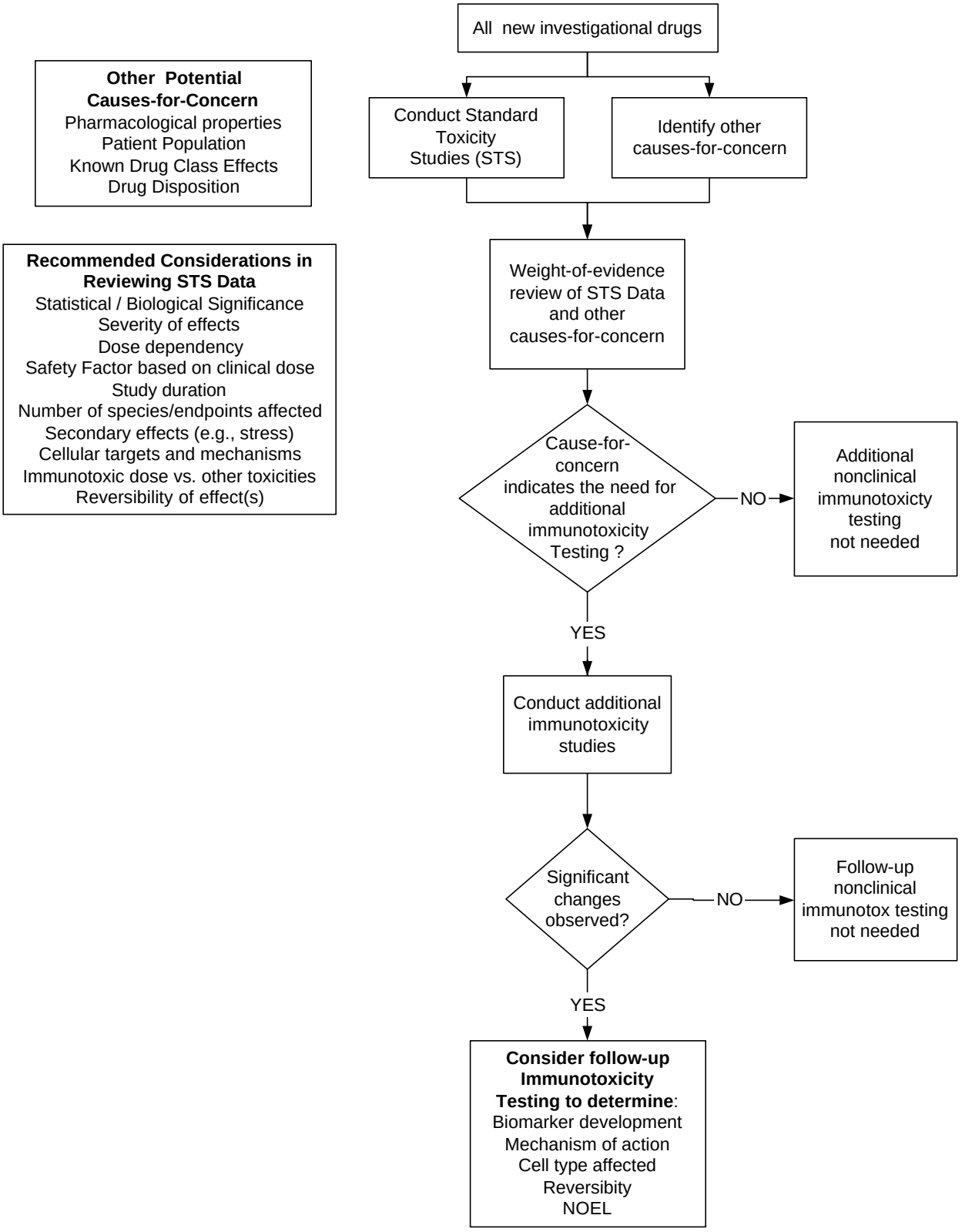
3. FOLLOW-UP IMMUNOTOXICITY STUDIES

If changes are observed with immunotoxicity testing, further studies should be considered to help determine the cell type affected and the mechanism of action. This type of information can provide more insight into potential risk and possibly lead to biomarker selection for clinical studies. These assays can include natural killer cell, host resistance or macrophage function assays. The findings from the STS and/or additional nonclinical immunotoxicity testing will help determine the need for, feasibility of, and type of clinical monitoring that is appropriate. In situations where the development candidate might have a pharmacological effect on the immune system, that specific component or associated function could be monitored. Additional guidance is beyond the scope of this guideline.

4. TIMING OF IMMUNOTOXICITY TESTING IN RELATION TO CLINICAL STUDIES

If the weight-of-evidence review indicates the need for additional immunotoxicity studies, these should be completed before exposure of a large population of patients to the drug. This will allow for the incorporation of immunotoxicity testing in the clinical studies if appropriate. The timing of the additional immunotoxicity testing might be determined by the nature of the effect by the test compound and the type of clinical testing that would be needed if a positive finding is observed with the additional immunotoxicity testing. If the target patient population is immunocompromised, immunotoxicity testing can be initiated at an earlier time point in the development of the drug.

Flow Diagram for Recommended Immunotoxicity Evaluation



Appendix 1

Methods to Evaluate Immunotoxicity

1. Standard Toxicity Studies

The following table lists the parameters that should be evaluated in standard toxicity studies for signs of immunotoxicity. These parameters (excluding hematology and serum chemistry) and methods in obtaining samples and evaluating tissue sections are described in more detail in documents from professional toxicological pathology societies.

Parameter	Specific Component
Hematology	Total leukocyte counts and absolute differential leukocyte counts
Clinical Chemistry	Globulin levels and A/G ratios
Gross pathology	Lymphoid organs / tissues
Organ weights	thymus, spleen, (optional: lymph nodes)
Histology	thymus, spleen, draining lymph node and at least one additional lymph node, bone marrow, Peyer's patch

1.1 Hematology and Clinical Chemistry

Total leukocyte counts (white blood cells) and absolute differential leukocyte counts are recommended to assess for immunotoxicity. When evaluating changes in globulin levels, other factors should be taken into account (e.g. nephrotoxicity). Changes in serum globulins can be an indication that there are changes in serum immunoglobulins. Although serum immunoglobulins are an insensitive indicator of immunosuppression, it can be useful in certain situations in order to better understand target cell populations or mechanism of action.

1.2 Gross Pathology and Organ Weights

All lymphoid tissues should be evaluated for gross changes at necropsy. However, this can be more difficult for the Peyer's patches of rats due to the small size. Spleen and thymus weights should be recorded. To minimize variability of spleen weights in dogs and monkeys, bleeding the animals thoroughly at necropsy is recommended. Atrophy of the thymus with ageing may preclude obtaining accurate thymus weight.

1.3 Histopathological Examination

Histopathological changes of the spleen and thymus should be evaluated as an indicator of systemic immunosuppression. The lymphoid tissue that drains or contacts the site of drug administration (and therefore is exposed to the highest concentration of the drug) should be examined. These sites include the Peyer's patches and mesenteric lymph nodes for orally administered drugs, bronchus-associated lymphoid tissues (BALT) for drugs administered by the inhalation route, nasal-associated lymphoid tissues (NALT) for drugs administered by the inhalation or nasal route, and the most proximal regional draining lymph nodes for drugs administered by the dermal, intramuscular, intradermal, intrathecal, or subcutaneous routes. The specific node selected should be at the discretion

of the sponsor based on the sponsors experience with the nodes. For intravenously administered drugs, the spleen can be considered the draining lymphoid tissue.

It is recommended that a "semi-quantitative" description of changes in compartments of lymphoid tissues should be used in recording changes and reporting treatment-related changes in lymphoid tissues.

1.4 Interpretation of Stress Related Changes

With standard toxicity studies, doses near or at the maximum tolerated dose can result in changes to the immune system related to stress. These effects on the immune system are most likely mediated by increased corticosterone or cortisol release. Commonly observed stress-related immune changes include increases in circulating neutrophils, decreases in circulating lymphocytes, decreases in thymus weight, decreases in thymic cortical cellularity and associated histopathologic changes ("starry sky" appearance), and changes in spleen and lymph node cellularity. Increases in adrenal gland weight can also be observed. In situations with clear clinical observations (eg. decrease body weight gain, decreased activity), some or all of the changes to lymphoid tissue and hematology parameters might be attributable to stress rather than to a direct immunotoxic effect. The evidence of stress should be compelling.

2. Additional Immunotoxicity Studies

2.1 Assay Characterization and Validation

In general, the immunotoxicity test selected should be widely used and have been demonstrated to be adequately sensitive and specific for known immunosuppressive agents. However, in certain situations, extensive validation might have not been completed and/or the assay might not be widely used. In these situations, a scientific / mechanistic basis for use of the assay is needed and if feasible, appropriate positive controls should be incorporated.

There can be variations of response for each type of immunotoxicity test used by different labs. In most situations, these changes do not affect the ability of the assay to assess immunotoxicity. However, to assure proper assay performance and lab proficiency, several standard technical validation parameters should be observed. These parameters can include determining intra- and inter-assay precision, technician-to-technician precision, limit of quantitation, linear region of quantitation and test sample stability. In addition, assay sensitivity to known immunosuppressive agents should be established. It is recommended that each laboratory conduct a positive control study periodically in order to demonstrate proficiency of performance, except for studies with non-human primates.

2.2 T-cell Dependent Antibody Response (TDAR)

The TDAR should be performed using a recognized T-cell dependent antigen (e.g. sheep red blood cells, SRBC or keyhole limpet hemocyanin, KLH) that result in a robust

antibody response. The endpoint selected should be shown to be the most appropriate for the chosen assay. For the SRBC assay, IgM is considered the most appropriate endpoint.

Antigens for immunization should not be used with adjuvants without justification. Alum might be acceptable for use only in non-human primate studies. The relative TDAR response can be strain-dependent, especially in mice. With outbred rats, there can be significant variability among rats within the same group. Inbred rat strains should not be used unless sufficient exposure data are provided.

Antibody can be measured by using an ELISA or other immunoassay methods. One advantage of this method over the antibody forming cell response is that samples can be collected serially during the study. In monkeys, serial blood collection can be important due to the high inter-animal variability in the kinetics of the response. For these studies, data may be expressed as the sum of the antibody response over several collection dates (eg. area under the curve).

ELISA results should be expressed either as concentration or titer, but expression as optical densities is not recommended.

When SRBC antigens are used for an ELISA, the preparation of the capture antigen that is coated on the plates is considered critical. SRBC capture antigen may be used as whole fixed erythrocytes or as membrane preparations.

For the rat TDAR and immunophenotyping assays, the addition of positive controls for each study with test compound might not be needed if the method used has been demonstrated to be adequately sensitive to immunosuppressive compounds.

2.3 Immunophenotyping

Immunophenotyping is the identification and/or enumeration of leukocyte subsets using antibodies. Immunophenotyping is usually conducted by flow cytometric analysis or by immunohistochemistry. With flow cytometry the percentage and absolute counts of a specific cell type or activation markers can be determined from large numbers of cells analyzed.

One of the advantages of immunohistochemistry over flow cytometry is that tissues from standard toxicity studies can be analyzed retrospectively if signs of immunotoxicity are observed. In addition, changes in cell types within a specific compartment within the lymphoid tissue can be observed. Some of the lymphocyte markers for certain species are sensitive to formalin fixation and can only be localized in tissue that are either fixed with certain fixatives or flash frozen. In addition, quantitation of leukocytes among subsets and intensity of staining is much more difficult with immunohistochemistry.

When immunophenotyping studies are used to characterize or identify alterations in specific leukocyte populations, the choice of the lymphoid organs to be evaluated should be based on changes observed. Immunophenotyping can be easily added to standard repeat dose toxicity studies and changes can be followed during the dosing phase and periods without drug exposure (reversal period). When flow cytometry is employed to enumerate specific cell populations, it is not a functional assay. However, flow cytometry can be used to measure antigen-specific immune responses of lymphocytes. Data obtained from peripheral blood can be useful as a bridge for clinical studies in which

peripheral blood leukocytes are also evaluated. It is recommended that absolute numbers of lymphocyte subsets rather than percentages be used in evaluating treatment-related changes.

2.4 Natural Killer Cell Activity Assays

Natural killer (NK) cell activity assays can be conducted if immunophenotyping studies demonstrate a decrease in number, or if STS studies demonstrate increased viral infection rates, or in response to other causes for concern. In general, all NK cell assays are ex vivo assay in which tissues (e.g. spleen) or blood are obtained from animals that have been treated with the test compound. Cell preparations are co-incubated with target cells that have been labeled with chromium. New methods that involve of non-radioactive labels can be used if adequately validated. Different effector to target cell ratios should be evaluated for each assay to obtain a sufficient level of cytotoxicity.

2.5 Host Resistance Studies

Host resistance studies involve challenging groups of mice or rats treated with the different doses of test compound with varying concentrations of a pathogen (bacteria, viral, parasitic) or tumor cells. Infectivity of the pathogens or tumor burden observed in vehicle versus test compound treated animals is used to determine if the test compound is able to alter host resistance. Models have been developed to evaluate a wide range of pathogens such as *Listeria monocytogenes*, *Streptoccous pneumoninae*, influenza virus, cytomegalovirus, *Plasmodium yoelii* and *Trichinella spiralis*. Tumor host resistance models in mice have used the B16F10 melanoma and PYB6 sarcoma tumor cell lines. Since the host defense mechanisms in these models are relatively well understood, it is recommended to choose a host resistance assay suitable for a specific target affected in standard toxicity studies or other immune toxicity tests.

Host resistance assays have an important role in identifying or confirming the cell type affected by a test compound. Since the host defense mechanisms against some microbes or tumor cells are well defined, test compound-related changes in these host resistance models would demonstrate that those cells types are targets. In addition, host resistance assay involve innate immune mechanisms in which specific immune function assays have not been developed. In conducting host resistance studies, the investigator should carefully consider the direct or indirect (non-immune mediated) effects of the test compound on the growth and pathogenicity of the organism or tumor cell. For instance, compounds that inhibit the proliferation of certain tumor cells can seem to increase host resistance.

2.6 Macrophage / Neutrophil Function

In vitro macrophage and neutrophil function assays (phagocytosis, oxidative burst) have been published for several species. These assays assess macrophage/neutrophil function of cells exposed to the test compound in vitro or obtained from animals treated with the test compound (ex vivo assay). In vitro exposure to test compound can also be investigated. An in vivo assay can also be used to assess the effects on the reticuloendothelial cell to phagocytize chromium-labeled sheep red blood cells. Animals

should be treated with the test compound and injected with chromium-labeled sheep red blood cells. Animals should be necropsied and tissues (eg. liver, spleen, lung, kidney) removed and radioactivity is counted.

2.7 Assays to Measure Cell-Mediated Immunity

Assays to measure cell-mediated immunity have not been as well established as those used for the antibody response. Delayed-type hypersensitivity (DTH) reactions with KLH immunization and challenge have been reported for mice. Cytotoxic T cell response can be generated in mice, however, since these involve the administration of a tumor cells line or graft, these assays are not considered toxicological studies. Reports on monkey DTH reactions have also been reported. However, these reactions in monkeys are very difficult to consistently reproduce. In addition, one should make sure that the DTH response is not mistaken for an antibody and complement mediated arthus reaction. Models in which contact sensitizers are used have been explored in mice but have not been well validated or extensively used.